Heritage and Sustainable Development ISSN 2712-0554
Vol. 1, No. 2, December 2019, pp.84-94

Brassica oleracea L. var. acephala ABA biosynthesis genes ($NCED_2$ and $NCED_3$) *in silico* interactome analysis

Ada Adilović[1], Ajdina Karić[2], Jasmin Šutković[3*]
[1,2,3] Genetic and Bioengineering, International University of Sarajevo, Bosnia

*Corresponding author: jsutkovic@ius.edu.ba

Abstract

Brassicaceae are industrially important flowering plants and in the same time often used for human consumption. Brassicae species are known heavy metal accumulators and therefore evaluated as potential phytoextraction plants. However, a detail analysis of heavy metal correlated genes, correlating to abscisic acid synthesis, is still not available. ABA (Abscisic acid) signaling genes from *Brassica oleracea*, $NCED_{2/3}$ are known to be correlated to heavy metal stresses, including cupper, arsenic and cadmium. $NCED_{2/3}$ proteins control the plants developmental processes and response to environmental stresses such as drought, heat stress, cold tolerance, and primarily, heavy metal influence. Through a simple *in silico* approach, we have confirmed that $NCED_{2/3}$ proteins share more than 65% similarity, as confirmed by phylogenetic analysis and 3D structure models. After the domain and interactome prediction, we revealed that $NCED_{2/3}$ strongly correlates with BoXDIX1, ABA_3, BoZE and BoZCD proteins, therefore indicating new roles to $NCED_{2/3}$ proteins, which is the involvement in the pigmentation and photoprotection processes of *Brassica oleracea.*

Keywords: Abscisic acid, Interactome, 3D model, *in silico*, *Brassica oleracea*

1 Introduction

Brassicaceae is a medium-sized and economically important family of flowering plants, informally known as mustard flowers, the crucifers, or the cabbage family. The family contains 372 genera and 4060 accepted species, such as *Brassica oleracea* (broccoli, cabbage, and cauliflower), *Brassica rapa* (turnip, Chinese cabbage), *Brassica napus* (rapeseed, etc.), *Arabidopsis thaliana* (thale cress) and many others [1].
Brassica species are used for human consumption, animal fodder, condiments, oil, and biofuel production, but also considered as a source of many nutrients such as carotenoids, tocopherols, different essential elements, carbohydrates and amino acids [2]. Besides, Brassicae species are known as metal accumulators and have been evaluated as potential phytoextraction plants, but detailed analysis of heavy metal correlated genes in Brassicae species is still not available.
Several studies have shown that heavy metal stress causes ABA (abscisic acid) correlated genes to be overexpressed, as shown in *Triticum aestivum*, *Solanum nigrum*, *Sedum alfredii* and *Pisum Sativum* [3,4,5,6,7]. Through a genome-wide classification and abiotic stress-responsive expression profiling, ABA (Abscisic acid) signaling genes, $NCED_2$, and $NCED_3$ were correlated to heavy metal stress in *Brassica oleracea* [8]. Also, a transcriptome analysis study showed a strong expression of ABA biosynthesis genes $NCED_2$ and $NCED_3$ in rice and under arsenic (As) stress [9]. In 2016, a study done on *Arabidopsis thaliana* showed that alteration of Cu homeostasis affected ABA biosynthesis, including several genes and among them $NCED_3$ [10]. Generally, (ABA) plant hormones have many roles in the plant developmental processes such as the growth of the organ size, seed and bud dormancy [11].
Further, it plays a vital role in response to environmental stresses such as drought, heat stress, cold tolerance, and primarily, heavy metal influence [12]. However, for further evaluation of the ABA correlated

biosynthesis genes, an *in-silico* research approach is a good alternative for a more detail analysis of NCED2 and NCED3, including their functional and structural characteristics.

2 Material and methods

2.1 Sequence retrieval and interactome analysis

The sequences for the NCED proteins ($NCED_2$ and $NCED_3$) were obtained from the National Center for Biotechnology Information (NCBI) database [13]. The accession numbers of the proteins can be found in Table 1.

Table 1. Accession numbers of *Brassica oleracea* $NCED_{2/3}$ proteins

PROTEIN NAME	SEQUENCE ACCESSION NUMBERS
$NCED_2$	106304727/ XP_013596575.1
$NCED_3$	106296147/ XP_013587670.1

To connect the NCED proteins with the proteins that have similar functions, we have done the interactome analysis. The Interactome was determined by using the STRING, a Protein-Protein interaction networks functional enrichment analysis online tool [14]. It was used for searching interlogs of the two proteins. The STRING database consists of known and predicted protein interactions of currently 24 million proteins and 5090 organisms. The program determines the binary interactions of each protein with predicted protein partners. Each interaction is then assigned to a confidence score which depicts the quality and number of an experimental technique used for the detection of these protein interactions [15]. The significance and reliability of the search was kept at high level by the confidence score of 0.7. For an additional and detailed genomic information of predicted proteins the Kyoto Encyclopedia of Genes and Genomes (KEGG) was accessed [16].

2.2 Multiple sequence alignment and domain prediction

After the interactome analysis and KEGG annotation, the multiple sequence alignment (MSA) analysis and domain prediction was performed.

For the Multiple sequence alignment (MSA) analysis, focusing on the variability and conserved patterns among the sequences of the proteins, we used Clustal W online software. Clustal W is available at the European Bioinformatics Institute (EMBL-EBI) website [17], with default options used. Clustal W is an MSA tool that uses seeded guide trees and HMM profile-profile techniques to deliver accurate alignments for any number of sequences [18,19].

Domains of the NCED proteins were identified by t the SMART (Simple Modular Architecture Research Tool) online tool, available on the website of EMBL (European Molecular Biology Laboratory). SMART also provides an analysis of domain architectures and detects more than 500 domain families in signaling, extracellular and chromatin-associated proteins [20, 21]. PFAM domain search was included [22] and the outlier homologues were also considered.

2.3 Phylogenetic tree construction

Following the MSA, a phylogenetic tree was constructed to visualize the similarity among the four proteins. Phylogeny.fr, a web service for the reconstruction of robust phylogenetic trees was used [23, 24]. Phylogeny.fr uses MUSCLE for alignment [25], Gblocks for curation [26], PhyML for phylogeny [28, 28] and Tree Dyn for tree rendering [29].

2.4 Localization of proteins

To check the subcellular localization of the proteins, we used the BUSCA online tool, an integrative web server to predict subcellular localization of proteins. BUSCA is able to predict 16 subcellular compartments for plant proteins, 9 subcellular compartments for animal or fungi proteins, 4 subcellular compartments for Gram-negative bacteria proteins and 3 subcellular compartments for Gram-positive bacteria proteins [30].

2.5 3D structure prediction

The prediction and visualization of the proteins were done in the Swiss-Model server, a fully automated protein structure homology-modeling server, allowing efficient 3D model quality check through an integrated Qmean scoring function, which is able to retrieve both local and global data-quality estimates on the basis of one single model. If the score is around zero, it indicates good agreement between the model structure and experimental structures of similar size. Scores of - 4.0 or below is an indication of models with low quality [31].

3 Results

3.1 Interactome of proteins

The predicted protein interaction is classified into physical or functional associations. Based on the STRING analysis, the query protein $NCED_2$ (XP_013596575.1) and $NCED_3$ (XP_013587670.1) resulted in 5 interactome proteins (see Figure 1 and Table 2). Further, through the additional analysis of uncharacterized proteins (106319391, 106295419, 106300040, Bo7g064130.1 and B05g021990.1) in KEGG online database, we correlated to specific proteins, as seen in Table 2.

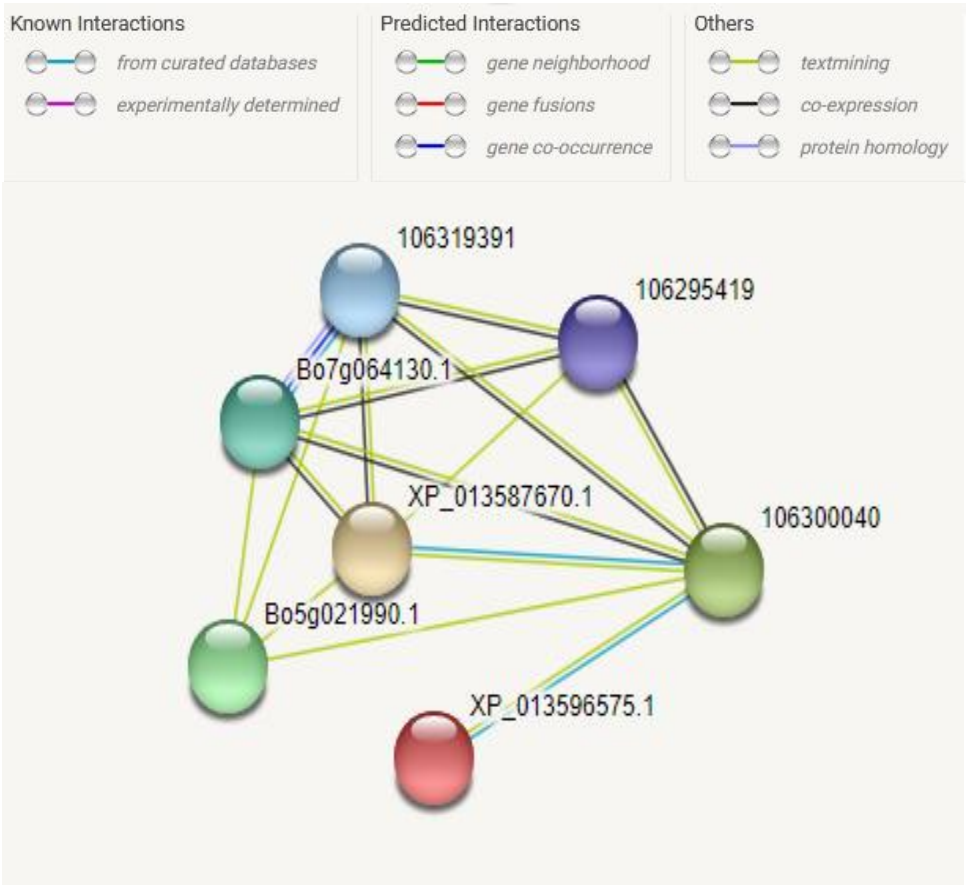

Figure 1. Interactome analysis by STRING of $NCED_2$ and $NCED_3$ (0.7 confidence score)

Table 2. Predicted Interactome partners and their functions base of KEGG annotation

No	*STRING/ NCBI IDs*	*Proposed Abbrev.*	*Score*	*Name/ function/Sequence link*
1	106300040/ XP_013591543	BoXDIX1	0.949	Xanthoxin dehydrogenase isoform X1, NAD(P)-binding Rossmann-fold superfamily protein (282 aa)

No	*STRING/ NCBI IDs*	*Proposed Abbrev.*	*Score*	*Name/ function/Sequence link*
2	Bo5g021990/ NP_564001.1	ABA3	0.832	Molybdenum cofactor sulfurase; Sulfurates the molybdenum cofactor (ABA3). Sulfation of molybdenum is essential for xanthine dehydrogenase (XDH) and aldehyde oxidase (ADO)(819 aa)
3	106319391/ XP_013613157	BoZE	0.805	Zeaxanthin epoxidase, chloroplastic; Converts zeaxanthin into antheraxanthin and subsequently violaxanthin (657 aa)
4	106295419/ XP_013586764	BoZCD	0.788	Zeta-carotene desaturase; Catalyzes the conversion of zeta-carotene to lycopene via the intermediary of neurosporene (561 aa)
5	Bo7g064130/ XP_013613157	N/A Protein	0.805	annotation not available (723 aa) – function unknown

Protein with the accession number Bo7g064130 will not be further analyzed due to the inability of functional annotation.

3.2 Multiple sequence alignment (MSA)

Table 3. Multiple Sequence Alignment scores (in %)

	BoZE	**BoZCD**	**ABA_3**	**BoXDIX1**	**$NCED_2$**	**$NCED_3$**
BoZE	100.00	12.01	19.15	15.95	15.87	16.43
BoZCD	12.01	100.00	16.98	13.14	15.49	16.84
ABA3	19.15	14.22	100.00	17.95	19.03	19.05
BoXDIX1	15.95	13.14	14.22	100.00	21.26	24.29
$NCED_2$	15.87	15.49	17.95	21.26	100.00	65.45
$NCED_3$	16.43	16.84	19.03	24.29	65.45	100.00

3.3 Domains of proteins

The domains in the two NCED proteins from *Brassica oleracea* were identified by the SMART software.

The results are presented in the table 4.

Table 4. Overview of main domains present in $NCED_2$ and $NCED_3$ proteins.

Protein name	***Domains***					
	RPE65		low complexity1		low complexity2	
	Start	End	Start	End	Start	End
NCED2	109	575	74	79	/	/
NCED3	127	589	25	37	92	100

The SMART analysis revealed that both NCED proteins have the same domains but on different locations. The RPE65 domain is a an enzyme Cartenoid, which is abundantly expressed in the retinal pigment epithelium ,where it catalyzed the formation of 11-cis-retinol from all-trans-retinyl esters. Carotenoids function as accessory photosynthetic pigments and as scavengers of oxygen radicals for photoprotection [32]. Low complexity regions are abundant regions, common in many protein sequences, but low information content. These regions show significant divergence across protein families and different genetic mechanisms, which resulted in degrees of compositional plasticity [33].

3.4 Phylogenetic tree

In Figure 2 we show the phylogenetic tree in a cladogram view, not revealing the common ancestors but the different geological time scale.

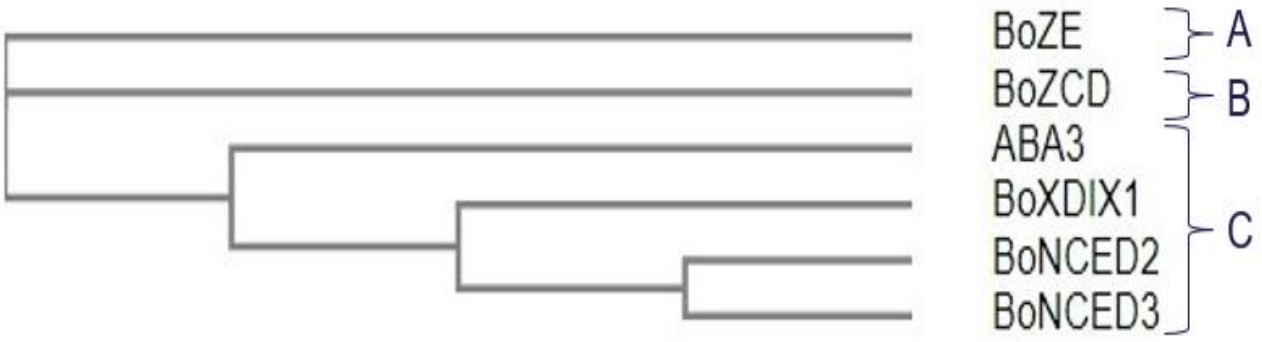

Figure 2. Phylogenetic tree (cladogram) of NCED proteins with other Brassicaceae family members

By constructing the phylogenetic tree, the evolutionary relations between the sequences are found. Based on the constructed Cladogram, we see 3 main clusters (A, B and C). Cluster A is represented by Zeaxanthin epoxidase (BoZE) protein, cluster B with Zeta-carotene desaturase (BoZCD) and the biggest cluster C includes the remaining 4 proteins; reorganized withing 3 subclusters. Cluster C closely correlates Abscisic acid $NCED_2$ and $NCED_3$ proteins with ABA3 and BoXDIX1 proteins.

3.5 Localization of proteins

BUSCA combines methods for identifying signal and transit peptides (DeepSig and TPpred3), GPI-anchors (PredGPI) and transmembrane domains (ENSEMBLE3.0 and BetAware) with tools for discriminating subcellular localization of both globular and membrane proteins (BaCelLo, MemLoci and SChloro) [30].

Results are shown for protein localization prediction in table below:

Table 5. Localization of $NCED_{2/3}$ proteins and their correlated interactomes

Protein ID	***Score***	***Localization class***
$NCED_2$	0.97	Chloroplast
$NCED_3$	0.97	Chloroplast
BoZE	0.94	Chloroplast
BoXDIX1	0.70	Chloroplast
BoZCD	0.76	Chloroplast membrane
ABA_3	0.76	Chloroplast

3.6 Predicted and verified 3D structure models

The 3D structure determination of proteins is essential for total understanding of the function, interactions, possible ligand prediction and conserved domain analysis. However, experimental determination of the 3D structure is a demanding and time-consuming process, so bioinformatics tools are used to predict the structures of proteins of interest. The 3D structures of NCED proteins predicted by Swiss-Model online server are seen in figure 3.

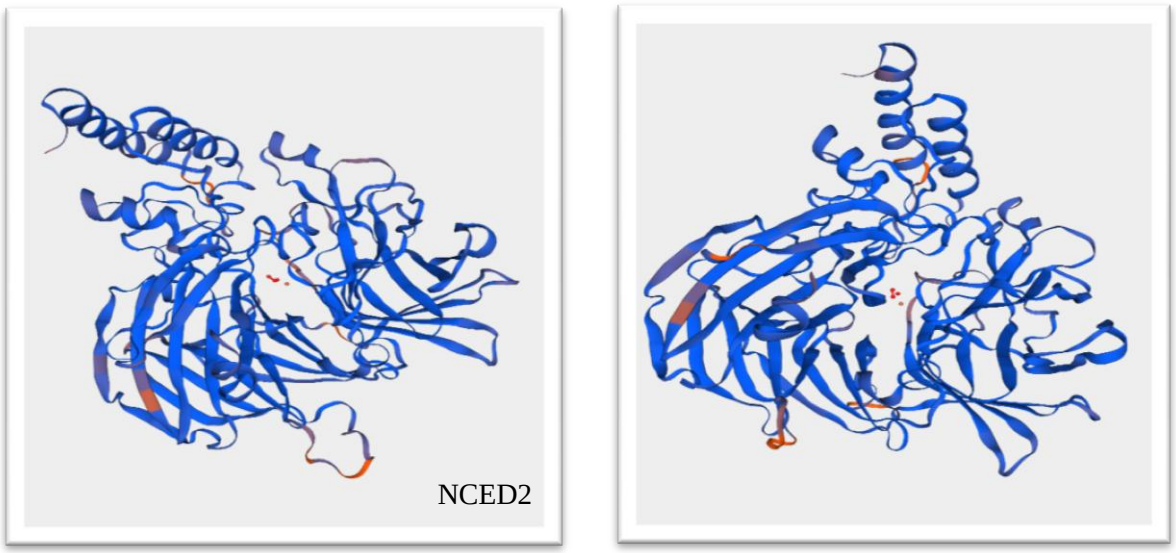

Figure 3. Predicted 3D structures of $NCED_2$ and $NCED_3$ proteins

As can be seen from the shown visualizations, both homologues have the same spatial order of specific domains, with a slight difference in structure shown mainly in terms of different looping regions and the number of helices and strands visualized.

In figure 4 we present the predicted 3D structures of the interactome partners for NCED2/3. Throughout the additional verification, as seen in table 6, we present the validation scored of all 4 proteins.

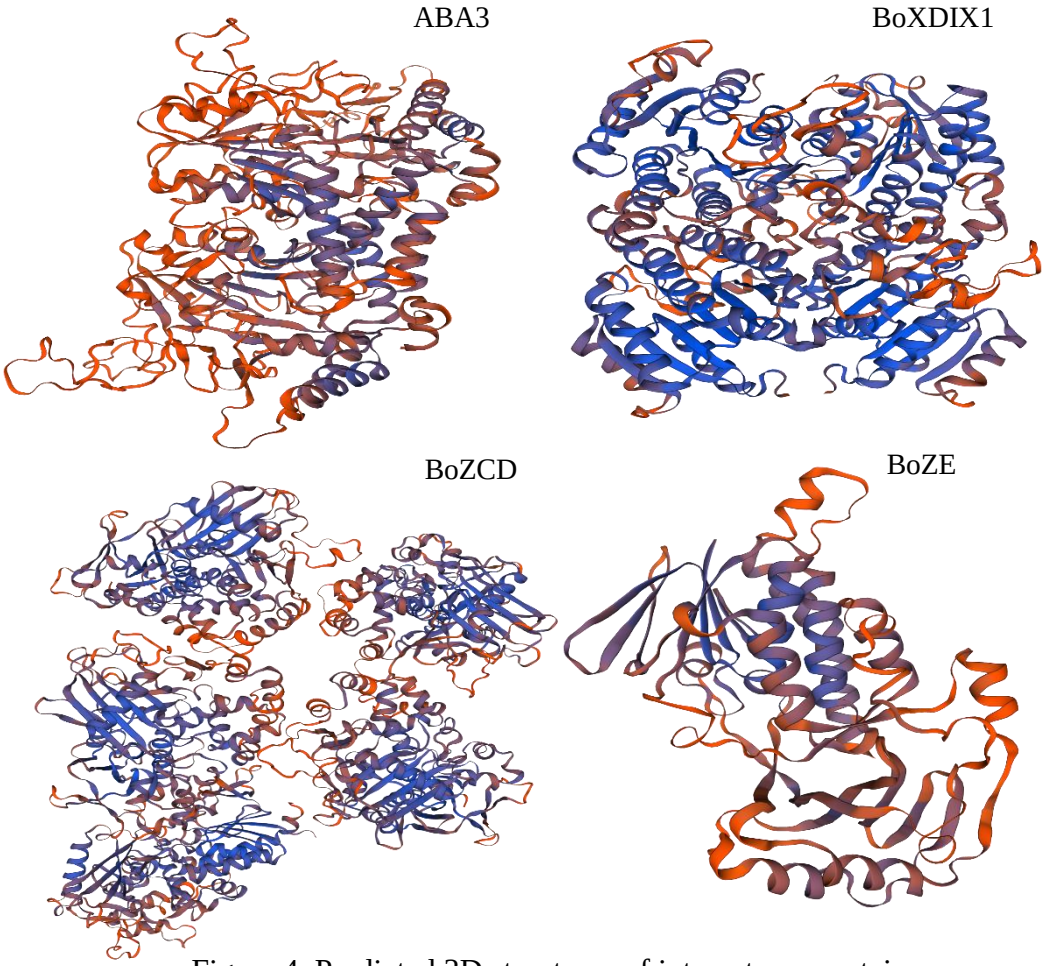

Figure 4. Predicted 3D structures of interactome proteins

As was already mentioned, the tools used were Qmean Z-score (If the score is around zero, it indicates good agreement between the model structure and experimental structures of similar size). Scores of -4.0 or below is an indication of models with low quality. Following the above described, the 3D structure prediction performed by SWISS MODEL server was sufficiently validated.

Table 6. Structure assessment via Qmean Z score

Name	***Qmean-Z score***
$NCED_2$	**-2.36**
$NCED_3$	**-1.47**
BoZE	**-3.25**
BoXDIX1	**-3.72**
BoZCD	**-4.05**
ABA_3	**-6.11**

4 Discussion and conclusion

Brasica oleracea is a known heavy metal accumulator plant, and as such a potential model for environmental phytoremediation processes. However, besides the positive attributions as phytoremediation plants, these metals-accumulating plants are directly or indirectly responsible for much of the dietary uptake of toxic heavy metals by humans and animals, causing potential danger for the consumers. Brassica family members, *cabbages,* are vegetables widely cultivated in wastewater-irrigated soils, so taking up heavy metals in large quantities to cause potential health risks to the consumers. Metal accumulation and translocation potential varies from plant to plant, and metal to metal [34]. In this regard, it is important to discover the potentials of metal-accumulator's viabilities of phytoremediator plants, is welcomed.

As mentioned above, few studies correlated ABA correlated genes to be overexpressed under heavy metal stress, as shown in *Triticum aestivum*, *Solanum nigrum*, *Sedum alfredii* and *Pisum Sativum* [3, 4, 5, 6, 7]. The ABA (Abscisic acid) signaling genes, $NCED_{2/3}$ were correlated to different heavy metal stress in *Brassica oleracea*, as under arsenic and cupper [8,9]. Besides heavy metal stress ABA genes are correlated to different plant developmental processes (growth of the organ size, seed and bud dormancy) and to environmental stresses such as drought, heat stress, cold tolerance, and primarily, heavy metal influence [11,12]. However, this is the first study to correlated $NCED_{2/3}$ with different interactome genes from *Brassica oleracea*, assigning the first 3D protein model to $NCED_{2/3}$ proteins. The predicted 3D model shared the same domain (RPE65) and cellular localization (Chloroplast).

In this work we have confirmed that $NCED_{2/3}$ proteins share strong homology, with 65.45% similarity, as seen in table 3. The phylogenetic tree analysis between the NCED proteins, additionally confirmed the similarity among these two proteins, where $NCED_2$ and $NCED_3$ share the same ancestor, separated by other groups in the tree. These results suggest that due to their close evolutionary relationship, they play important biochemical roles by performing same or similar functions within the cell.

After the interactome analysis, we revealed that NCED2 and NCED3 strongly correlated with BoXDIX1, ABA3, BoZE and BoZCD proteins (0.949, 0.832, 0.805 and 0.788). Before the analysis, in order to limit a higher number of interactor proteins, the confidence level was set to minimum of 0.7 and to show maximum 10 interactors.

BoXDIX1, a xanthoxin dehydrogenase protein, is shown to be correlated the biosynthesis of thiamine [35] and ABA [36], respectively. In addition, xanthoxin dehydrogenase is reported to participate in proline biosynthesis and the sugar-mediated signaling pathway [37]. According the literature review, this strong interaction between BoXDIX1 and $NCED_{2/3}$ was expected.

The ABA correlated genes are known for their role in abiotic stresses [38]. In 2001, in *A.thaliana* the ABA_3 is presented as a molybdenum cofactor sulfurase required for activation of aldehyde oxidase and xanthine dehydrogenase [39]. *Brassica oleracea var. acephala* is a purple ornamental cabbage, also popular as a decorative plant, cultivated for its colorful leaf rosettes that persist in cool weather. In 2019, a study showed that *B.oleracea* genes ($BoNCED_{2.1}$, $BoNCED_{2.2}$, $BoNCED_6$, $BoNCED_{9.1}$,and $BoAAO_{3.2}$) were significantly higher in purple compared to green leaves, including ABA_3 [40]. Due to the strong interaction with, both NCED proteins, we suggest a possible role of $NCED_3$ in the pigmentation of the inner leaves of purple ornamental cabbage, not shown in available literature reviews.

According several studies, the BoZe (Zeaxanthin epoxidase) protein is confirmed to control the ABA biosynthesis pathways [41, 42, 43]. However, recent studies showed that Zeaxanthin epoxidase proteins also confer drought and salt tolerance in plant, like transgenic tobacco and tomato [44]. Further, DtZEP gene expression responded rapidly to light irradiation and hyperosmotic stress [45], as shown and confirmed in NCED genes [11, 12].
The weakest interaction of NCED proteins is show with a protein of a Carotenoid family, a Zeta-carotene desaturase protein (annotated as BoZCD). These proteins are pigments synthetized by plants, fungi, bacteria, and algae with the main function of protecting them from the action of singlet oxygen and other radicals [46]. Carotenoids control various physiological processes, including growth, development, and environmental affect control. In green tissues, carotenoids act as accessory pigments for the assembly of photosystems and light-harvesting antenna complexes and also have photo-protective functions during photosynthesis [47, 48, 49]. According the current literature overview, $NCED_{2/3}$ role was not connected to plant pigmentation processes.
However, according the domain prediction result, the RPE65 domain predicted in both NCED protein, belongs to the Cartenoid enzyme family, abundantly expressed in the retinal pigment epithelium , catalyzing the formation of 11-cis-retinol from all-trans-retinyl esters. Accordingly, the Carotenoids function as accessory photosynthetic pigments and as scavengers of oxygen radicals for photoprotection [32]. The obtained in *silico* domain and interactome prediction results suggests $NCED_{2/3}$ role as an accessory pigment for the assembly of photosystem and photoprotection.
The 3D structures of proteins enable additional functional studies, domain analysis, molecular interaction studies, estimation of structural similarity between proteins etc. In this study, we used Swiss-Model online tool, a protein homology modeling server, used to create models of target proteins. These models contain information about the tendency for mutation of each amino acid in a sequence and are unique for each protein. They are created for a set of known 3D structures as well as for the user sequence, and then scanned to find a match [31].
Further, Swiss-Model server, includes Qmean verification tool, enabling additional confirmation and verification of the modeled structures, in the form of Z-score. The Z-score analysis showed that NCED models have negative Z-scores (in average of -2 Z-score), an accepted score for further structural validations. The verification results of both NCED proteins in *B.oleracea*, showed sufficient quality, required for further analysis (-2.36 for $NCED_2$ and -1.47 for $NCED_3$). Models of low quality are expected to have strongly negative Qmean Z-scores, less than -3.5 [31]. However, BoZE and BoXDIX1 interactome proteins showed lower model quality than expected (-3.25, -3.72) and BoZCD and ABA_3 showed scores of unaccepted values, -4/05 and – 6.11, respectively.

In conclusion, it is through bioinformatics analysis that we identified and structurally predicted the $NCED_{2/3}$ protein structures and their interactome models. $NCED_{2/3}$ proteins have shown significant similarity in protein sequence and 3D structure. The phylogenetic analysis revealed similar evolutionary paths of evolvement, these 2 proteins sharing common ancestor with BoXDIX1, a close related protein, sharing same functions. The results obtained in this study lead us to the conclusion that cellular functions of the 2 homologues NCED proteins are very similar, whereas their interactomes have shown similar and diverse functions. However, two partner interactome proteins, ABA3 and BoZCD, involved in pigmentation of the inner leaves and in the pigmentation process of the photosystem. Therefore, due to a strong interaction to ABA3 and BoZCD proteins, we may suggest a novel function to $NCED_{2/3}$ proteins, in the pigmentation mechanism of plant photosystem. Experimental determination of 3D structures of all analyzed proteins is required, as well as further testing in terms of interactome and co-localization analysis, to completely understand the role of NCED proteins in various physiological processes.

5 References

[1] I. Ahuja, J. Rohloff, and A. M. Bones, “Defence mechanisms of Brassicaceae: Implications for plant-insect interactions and potential for integrated pest management. A review,” *Agronomy for Sustainable Development*, vol. 30, no. 2. Springer, pp. 311–348, 30-Apr-2010, doi: 10.1051/agro/2009025.

[2] M. P. Mourato, I. N. Moreira, I. Leitão, F. R. Pinto, J. R. Sales, and L. L. Martins, “Effect of Heavy

Metals in Plants of the Genus Brassica," *Int. J. Mol. Sci*, vol. 16, pp. 17975–17998, 2015, doi: 10.3390/ijms160817975.

[3] F. M. Shakirova, M. V. Bezrukova, C. R. Allagulova, D. R. Maslennikova, and A. R. Lubyanova, "Wheat germ agglutinin and dehydrins as ABA-regulated components of SA-induced cadmium resistance in wheat plants," in *Salicylic Acid: A Multifaceted Hormone*, Springer Singapore, 2017, pp. 77–96.

[4] G. B. Pompeu *et al.*, "Abscisic acid-deficient sit tomato mutant responses to cadmium-induced stress," *Springer*, doi: 10.1007/s00709-016-0989-4.

[5] Q. Lu, G. Sun, J. Liu, and Y. Tang, "Effects of Abscisic Acid on Growth and Cadmium Accumulation of Pea Seedlings," 2018.

[6] Q. Lu, S. Chen, Y. Li, F. Zheng, B. He, and M. Gu, "Exogenous abscisic acid (ABA) promotes cadmium (Cd) accumulation in Sedum alfredii Hance by regulating the expression of Cd stress response genes," *Environ. Sci. Pollut. Res.*, vol. 27, no. 8, pp. 8719–8731, Mar. 2020, doi: 10.1007/s11356-019-07512-w.

[7] J. Wang *et al.*, "The effects of abscisic acid (ABA) addition on cadmium accumulation of two ecotypes of Solanum photeinocarpum," *Environ. Monit. Assess.*, vol. 188, no. 3, pp. 1–8, Mar. 2016, doi: 10.1007/s10661-016-5194-6.

[8] Y. Kim, I. Hwang, H. Jung, J. Park, … J. K.-J. of plant growth, and undefined 2016, "Genome-Wide Classification and Abiotic Stress-Responsive Expression Profiling of Carotenoid Oxygenase Genes in Brassica rapa and Brassica oleracea," *Springer*.

[9] T.-L. Huang *et al.*, "Transcriptomic changes and signalling pathways induced by arsenic stress in rice roots Elongation factor 1 gamma," *Plant Mol Biol*, doi: 10.1007/s11103-012-9969-z.

[10] À. Carrió-Seguí, P. Romero, … A. S.-P. and C., and undefined 2016, "Interaction Between ABA Signaling and Copper Homeostasis in Arabidopsis thaliana," *academic.oup.com*.

[11] L. Bücker-Neto, A. Paiva, R. M.-… and molecular biology, and undefined 2017, "Interactions between plant hormones and heavy metals responses," *SciELO Bras.*

[12] T. Kuromori, M. Seo, K. Shinozaki, ABA transport and plant water stress responses. Trends in plant science, vol. 23, no. 6, pp. 513-522, 2018.

[13] E. W. Sayers *et al.*, "Database resources of the national center for biotechnology information," *Nucleic Acids Res.*, vol. 39, no. SUPPL. 1, pp. D38–D51, Jan. 2011, doi: 10.1093/nar/gkq1172.

[14] L. J. Jensen *et al.*, "STRING 8 - A global view on proteins and their functional interactions in 630 organisms," *Nucleic Acids Res.*, vol. 37, no. SUPPL. 1, pp. D412–D416, Jan. 2009, doi: 10.1093/nar/gkn760.

[15] D. Szklarczyk *et al.*, "STRING v10: Protein-protein interaction networks, integrated over the tree of life," *Nucleic Acids Res.*, vol. 43, no. D1, pp. D447–D452, 2018.

[16] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research, vol. 28, no. 1, pp. 27-30, 2000.

[17] H. Fu, J. H. Doelling,C. S. Arendt, M. Hochstrasser, R. D. Vierstra, Molecular organization of the 20S proteasome gene family from aArabidopsis thaliana. Genetics, vol. 149, no. 2, pp. 677-692, 1998.

[18] M. Goujon, H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, R. Lopez, A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res., vol. 38, W695- W699, 2010.

[19] F. Sievers, A. Wilm, D. G. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, D. Higgins, Fast, scalable generation of high quality protein multiple sequence alignments using Clustal Omega.Mol. Syst. Biol., vol. 7, 539, 2011.

[20] J. Schultz, F. Milpetz, P. Bork, C. P. Ponting, SMART, a simple modular architecture research tool: Identification of signaling domains. Proc. Natl. Acad. Sci. USA, vol. 95, pp. 5857-5864, 1998.

[21] I. Letunic, T. Doerks, P. Bork, SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res., vol. 40, pp. 302-305, 2012.

[22] M. Punta, P. C. Coggill, R. Y. Eberhardt, J. Mistry, J. Tate, C. Boursnell, N. Pang, K. Forslund, G. Ceric, J. Clements, A. Heger, L. Holm, E. L. L. Sonnhammer, S. R. Eddy, A. Bateman, R. D. Finn, The Pfam protein families database. Nucleic Acids Res., vol. 42, pp. 222-230, 2014.

[23] A. Dereeper, S. Audic, J. M. Claverie, G. Blanc, BLASTEXPLORER helps you building datasets for phylogenetic analysis. BMC Evol. Biol., vol.10, 8, 2010.

[24] A. Dereeper, V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet, J. F. Dufayard, S. Guindon, V. Lefort, M. Lescot, J. M. Claverie, O. Gascuel, Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res., vol. 36, pp. 465-469, 2008.

[25] P. V. Eijk, Nucleotide excision repair in yeast. PhD thesis, Leiden University, 2012.

[26] J. Castresana, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. Vol. 17, no. 4, pp. 540-552, 2012.

[27] S. Guindon, O. Gascuel, A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol., vol. 52, no. 5, pp. 696-704, 2003.

[28] M. Anisimova, O. Gascuel, Approximate likelihood ratio test for branches: A fast, accurate and powerful alternative. Syst. Biol., vol. 55, no. 4, pp. 539-552, 2006.

[29] F. Chevenet, C. Brun, A. L. Banuls, B. Jacq, R. Chisten, TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics, vol. 7, 439, 2006.

[30] C. Savojardo, P. L. Martelli, P. Fariselli, G. Profiti, R. Casadio, BUSCA: an integrative web server to predict subcellular localization of proteins. Nucleid Acids Research, vol. 46, pp. 459-466, 2018.

[31] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli, T. Schwede, SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res., vol. 42, pp. 252-258, 2014.

[32] Wyss, Adrian. Carotene oxygenases: a new family of double bond cleavage enzymes. The Journal of Nutrition, vol. 134, no. 1, pp. 246-250, 2004.

[33] A. Coletta, J. W. Pinney, D. Y. W. Solís, J. Marsh, S. R. Pettifer, and T. K. Attwood, "Low-complexity regions within protein sequences have position-dependent roles," *BMC Syst. Biol.*, vol. 4, Apr. 2010, doi: 10.1186/1752-0509-4-43.

[34] G. A. Boamponsem, M. Kumi, I. Debrah. "Heavy metals accumulation in cabbage, lettuce and carrot irrigated with wastewater from Nagodi mining site in Ghana." Int J Sci Technol Res vol. 1, no. 11, pp. 124-129, 2012.

[35] M. Arefian, S. Vessal, S. Malekzadeh-Shafaroudi, K. H. Siddique, A. Bagheri, Comparative proteomics and gene expression analyses revealed responsive proteins and mechanisms for salt tolerance in chickpea genotypes. BMC plant biology, vol. 19, no. 1, pp. 300, 2019.

[36] S. H. Schwartz, K. M. Leon-Kloosterziel, M. Koornneef, J. A. Zeevaart, Biochemical characterization of the ABA2 and ABA3 mutants in Arabidopsis thaliana. Plant physiology, vol. 114, no. 1, pp. 161-166, 1997.

[37] P. León, J. Sheen. "Sugar and hormone connections." Trends in plant science vol. 8, no. 3, pp. 110-116, 2003.

[38] Y. Lu *et al.*, "Overexpression of Arabidopsis Molybdenum Cofactor Sulfurase Gene Confers Drought Tolerance in Maize (Zea mays L.)," *PLoS One*, vol. 8, no. 1, Jan. 2013, doi: 10.1371/journal.pone.0052126.

[39] F. Bittner, M. Oreb, and R. R. Mendel, "ABA3 Is a Molybdenum Cofactor Sulfurase Required for Activation of Aldehyde Oxidase and Xanthine Dehydrogenase in Arabidopsis thaliana*," *ASBMB*, 2001, doi: 10.1074/jbc.C100472200.

[40] S. W. Jin *et al.*, "Abscisic acid and ethylene biosynthesis-related genes are associated with

anthocyanin accumulation in purple ornamental cabbage (Brassica oleracea var. Acephala)," *Genome*, vol. 62, no. 8, pp. 513–526, 2019, doi: 10.1139/gen-2019-0038.

[41] C. Audran, C., *et al.*, Expression studies of the zeaxanthin epoxidase gene inNicotiana plumbaginifolia. Plant Physiology, vol. 118, no. 3, pp. 1021-1028, 1998.

[42] Z. Zhang, *et al.*, MsZEP, a novel zeaxanthin epoxidase gene from alfalfa (Medicago sativa), confers drought and salt tolerance in transgenic tobacco. Plant cell reports, vol. 35, no. 2, pp. 439-453, 2016.

[43] E. Marin *et al.*, "Molecular identification of zeaxanthin epoxidase of Nicotiana plumbaginifolia, a gene involved in abscisic acid biosynthesis and corresponding to the ABA locus of Arabidopsis thaliana," *EMBO J.*, vol. 15, no. 10, pp. 2331–2342, 1996, doi: 10.1002/j.1460-2075.1996.tb00589.x.

[44] Y. Cao, Z. Zhang, T. Zhang, J. An, … L. C.-B., and undefined 2019, "Overexpression of the alfalfa zeaxanthin epoxidase gene delays seed germination in transgenic tobacco," *bp.ueb.cas.cz*, doi: 10.32615/bp.2019.059.

[45] M. Kim, Y. Kang, and E. Jin, "Gene Expression Analysis of Zeaxanthin Epoxidase from the Marine Microalga Dunaliella tertiolecta in Response to Light/Dark Cycle and Salinity," *J. Microbiol. Biotechnol*, vol. 29, no. 9, pp. 1453–1459, 2019, doi: 10.4014/jmb.1904.04053.

[46] H. Sies, W. Stahl, A. S.-A. of the N. Y. A. of, and undefined 1992, "Antioxidant Functions of Vitamins: Vitamins E and C, Beta- Carotene, and Other Carotenoidsa," *academia.edu*, Accessed: Jan. 23, 2021. [Online]. Available: https://www.academia.edu/download/46050134/j.1749-6632.1992.tb17085.x20160529-12918-r9vk38.pdf.

[47] H. Dong, et al. The Arabidopsis spontaneous cell death1 gene, encoding a zeta-carotene desaturase essential for carotenoid biosynthesis, is involved in chloroplast development, photoprotection and retrograde signalling. Cell Res., vol. 17, pp. 458–470, 2007.

[48] L. Pizarro and C. Stange, "Light-dependent regulation of carotenoid biosynthesis in plants," *Ciencia e Investigacion Agraria*, vol. 36, no. 2. Pontificia Universidad Catolica de Chile, Facultad de Agronomia e Ingenieria Forestal, pp. 143–162, 2009, doi: 10.4067/S0718-16202009000200001.

[49] N. Nisar, L. Li, S. Lu, N. C. Khin, and B. J. Pogson, "Carotenoid metabolism in plants," *Molecular Plant*, vol. 8, no. 1. Cell Press, pp. 68–82, Jan. 05, 2015, doi: 10.1016/j.molp.2014.12.007.